Circular No. 816

August 1949 Washington, D.C.

UNITED STATES DEPARTMENT OF AGRICULTURE

Grasshopper Egg-Pod Distribution in the Northern Great Plains and Its Relation to Egg-Survey Methods

By E. G. Davis, *entomologist, Division of Cereal and Forage Insect Investigations*, and F. M. Wadley, *statistical consultant, Bureau of Entomology and Plant Quarantine, Agricultural Research Administration*[1]

CONTENTS

[1]The authors acknowledge the assistance given by J. R. Parker, in charge of the Bozeman, Mont., laboratory, Division of Cereal and Forage Insect Investigations, in both the field work and the preparation of the manuscript, and by C. M. Packard, in charge of the Division, for helpful criticism. Members of the staffs of the Montana State entomologist's office, the Division of Grasshopper Control, and the Bozeman laboratory aided in much of the field work.

INTRODUCTION

Grasshopper surveys supply the preliminary information needed for the most economical and successful conduct of control campaigns. This basic information consists of (1) the location and extent of the infestation, and (2) its density. An important consideration pertaining to grasshopper surveys is their timeliness. The most successful control campaigns are necessarily planned well in advance of the field work, and surveys made early enough must maintain their reliability up to the time field control work begins. An egg survey made in the fall of the year gives the best results.

For at least 25 years grasshopper surveys have been employed to predict future populations. During that time the survey has evolved from one of reconnaissance to the present standardized type in which infestations are classified into categories based on the number of egg pods per square foot of soil sampled.[2 3] Fall egg surveys are conducted by the Division of Grasshopper Control throughout the area subject to severe grasshopper damage, which includes nearly all the States west of the Mississippi River. Usually, however, it is necessary to survey only a portion of this area in any one year. It has been customary to sample an average of 7 to 11 fields representative of each county, to obtain mean egg-pod populations.

The egg survey is yielding serviceable results. In its present form, however, it requires a fairly large personnel and a great deal of travel. Delays due to adverse weather during the period available for surveying add to the labor problem. Any improvement that could reduce, without loss of accuracy, either the labor or the travel involved would result in a considerable saving of money. It also seems probable that some gain in accuracy can be obtained. Although it is evident that nothing can be done about the size of the area to be surveyed or about the prevailing weather, it is believed that field and county sampling methods can be improved.

The conclusions drawn from the data presented in this paper have been found to apply to conditions prevailing in the northern Great Plains, where the species *Melanoplus mexicanus* (Sauss.) was dominant while the data were being obtained. Whether they will apply to other areas or species has not been determined; however, many of the principles of sampling will be applicable. The present paper is on the application of these and other related studies to problems of field population and sampling, especially to immediate problems of survey.

SOURCES OF DATA

Intensive studies of grasshoppers in typical environments were initiated by the Bureau's grasshopper research station at Bozeman, Mont., in 1931. The studies were conducted on eight areas, each approximately 2

---

[2] Shotwell, R. L. A METHOD FOR MAKING A GRASSHOPPER SURVEY. Jour. Econ. Ent. 28: 486-491. 1935.

[3] ——— SOME PROBLEMS OF THE ANNUAL GRASSHOPPER SURVEY. Jour. Econ. Ent. 31: 523-533. 1938.

miles wide and 4 miles long, in Montana, North Dakota, and South Dakota. One of the main objectives was to test and improve grasshopper-survey methods.

In the egg surveys made during the earlier years of the studies, 10 or more 1-square-foot samples were taken on every quarter section of land. By 1938 doubt had arisen as to whether the number of samples taken on the areas was too few or too many, and whether taking a fixed number of samples per unit area is the most accurate way of determining mean egg-pod populations in the various crops and habitats. That the personnel charged with conducting the general surveys were also in doubt regarding the adequacy of sampling is evident from the changes made in the sampling instructions given the surveyors. To determine the dependability of sampling procedure several counties were surveyed in sufficient detail to obtain data for statistical analysis.

The first special intensive egg survey was conducted in the fall of 1939 in Fergus and Judith Basin Counties, in north-central Montana. With the aid of one member of the general survey staff, 200 fields were examined in 2 counties, which were treated as a single unit. In 1940, with the assistance of survey personnel supplied by the Division of Grasshopper Control and the Montana State entomologist, special egg surveys were conducted in 10 counties in north-central Montana, 70 fields in each county being sampled. The last of the special egg surveys was conducted in Davison and Brown Counties, South Dakota, in 1942. After the special county egg surveys were completed, the data from them and from the study areas were analyzed by methods developed along the lines shown by Snedecor.[4]

EGG-POD DISTRIBUTION

In Relation to Types of Vegetation Habitats

In any adequate study of sampling methods the distribution of the egg pods among the different habitats must be considered. Although it has been common knowledge for years that grasshoppers are selective in their egg-laying habits, accurate comparisons of the egg populations in habitats were lacking. Intensive studies were made, therefore, of egg-pod distribution in fields and field margins.

Records of egg populations in different plant associations on the northern Great Plains study areas for the 8-year period 1936-43 are summarized in table 1. It can readily be seen that some habitats, such as field margins and idle land, contain greater egg concentrations than others. As to egg-pod category, the first four habitats listed could be classed as high, the next four as medium, and the last five as low. Plowed and fallowed lands, not listed in the table, had practically no egg pods. Although the position of each habitat varied somewhat from year to year, as compared with its 8-year average, most of the variations were not great enough to change a habitat from one category to another. This tendency to remain in about the same relative position from year to year led to the idea that a single habitat might be used in carrying on the general survey.

[4]Snedecor, G. W. statistical methods, Ed. 3, 422 pp., Ames, Iowa. 1940.

TABLE 1.—*Average egg-pod distribution per square foot by habitats for areas under study in the northern Great Plains, 1936–43*

Habitat	1936	1937	1938	1939	1940	1941	1942	1943	Average 1936-43
Field margins	0.761	0.741	1.241	0.725	0.862	0.773	0.976	0.816	0.862
Idle land	.581	.394	.654	.624	.519	.615	.553	.440	.547
Pasture	.400	.240	.575	1.000	.301	.300	.095	.385	.412
Legumes	.278	.514	.421	.574	.418	.353	.355	.314	.403
Flax stubble	......	......	......	......	.101	.718	.267	.317	.351
Small-grain stubble	.336	.546	.356	.587	.249	.259	.279	.162	.347
Prairie	.526	.387	.369	.305	.389	.293	.120	.169	.320
Crested wheatgrass	.375	.600	.580	.280	.140	.175	.182	.196	.316
Wild-hay meadows	.000	.250	.730	.090	.214	.280	.040	.000	.200
River-bottom land	.050	.000	.330	.000	.330	.120	.200	.120	.144
Truck crops	.390	.207	.133	.210	.006	.107	.000	.013	.133
Corn	.165	.174	.180	.104	.134	.190	.032	.023	.125
Sorghums	......	......	......	.185	.249	......	.027	.000	.115

The data in table 1 are quantitative only to the extent that they indicate the egg-pod density in the respective habitats. To apply these data to general field conditions, the percentage of the total area occupied by each habitat must be considered. For example, although of all the habitats field margins contain the heaviest populations, they comprise only about 5 percent of the total area. Conversely, small-grain fields, which have moderate populations, comprise about 80 percent of the farmed area. In the 10-county special survey, grainfields predominated and crop differences were not prominent; hence, crop-type differences were not considered in further analysis of these data, except for distinction between field and margin.

In Fields and Margins

One of the first essentials in the development of a dependable survey method is a knowledge of egg-pod distribution within fields. One opinion was that more eggs were laid near the edge of the field than near the center. To obtain more definite data, egg sampling was conducted in 10 counties in north-central Montana, samples being taken from 70 fields located at random in each county. The number of fields of each major crop sampled was in proportion to the acreage of that crop in the county. In each field 5 pairs of ½-square foot units of soil were taken at equal distances apart in a straight line from the edge of the field to the center. Ten ½-square-foot units were also taken in the uncultivated margin of each field, 5 units in the half nearest the crop, and the other 5 in the half nearest the road.

The mean egg-pod population for each field and margin location is shown by counties in table 2. It will be noted that populations for the different locations within fields tended to vary at random. In the margins numbers tended to be greater in the half near the field than in the half near the road.

TABLE 2.—*Mean egg-pod populations per square foot in fields and margins in 10 counties in north-central Montana in 1940. (Five field and 2 marginal locations were sampled at each of 70 different places in each county)*

County	Field locations[1]					Marginal locations	
	1	2	3	4	5	Near road	Near field
Blaine	0.80	0.33	0.56	0.63	0.63	0.75	0.89
Cascade	.48	.44	.64	.86	.30	.56	.72
Chouteau	.64	.64	.36	.33	.34	.79	1.00
Fergus	.62	.30	.32	.38	.46	.48	.84
Hill	.20	.26	.42	.50	.64	.82	1.56
Judith Basin	.14	.04	.10	.19	.04	.06	.10
Liberty	.34	.23	.46	.43	.36	.55	.80
Pondera	.56	.79	.66	.96	.89	2.50	2.91
Teton	.59	.67	.49	.49	.63	1.85	1.90
Toole	.34	.34	.53	.39	.44	.43	.58
Average	.47	.40	.45	.52	.47	.88	1.13

[1]Locations designated by numbers: (1) Edge of field; (2) one-fourth of the distance to center of field; (3) one-half of the distance to center of field; (4) three-fourths of the distance to center of field; (5) center of field.

To obtain further information on the effect of sample location, the data were studied by an analysis of variance. The results are shown in table 3. The arrangement of the sampling makes it possible to study the

TABLE 3.—*Variance in numbers of egg pods per square foot between fields, between locations within fields, and between units within locations, for 10 counties in north-central Montana in 1940 (degrees of freedom given in parentheses in column heading)*

County	Between fields (69)	Between locations (4)	Interaction, field and location (276)	Units within[1] locations (350)
Blaine	[2]1.01	[3]1.02	0.33	0.42
Cascade	[2].93	[3]1.60	[3].50	.35
Chouteau	[2].80	[3].93	.31	.36
Fergus	[3].50	.63	.30	.32
Hill	[2]1.07	1.03	.45	.42
Judith Basin	[3].09	[3].14	.05	.07
Liberty	[2].37	.28	.24	.16
Pondera	[2].60	.92	.42	.38
Teton	[2]1.44	.25	[2].82	.47
Toole	[2].68	.22	.18	.24

[1]Interaction, *field and location*, is typically error for *location; units* is error for *interaction* and for fields in comparing variances by the *F* test.
[2]Significant at 1-percent level of probability.
[3]Significant at 5-percent level of probability.

over-all effect of given locations in the field, as well as the interaction or differential effect of location from field to field. The variation between units, as shown within pairs at each location, can also be studied. As

might be expected, there was a significant difference between field populations. In all but 2 counties (Fergus and Judith Basin) the differences were highly significant. The differences between populations in the 5 different within-field locations, however, showed a slight tendency toward significance in only 4 of the 10 counties, as seems apparent from a comparison of the averages in table 2.

It should make no material difference, therefore, where the units are taken within fields if they are fairly well distributed. Some data of this type might be considered to need transformation before analysis, but population levels were rather uniform, and the total number of egg pods found in each designation was high enough so that little or no gain could be expected from the transformation.

The results of an analysis of the data for the field margins are given in table 4. These results are similar to those for the fields. Population

Table 4.—*Variance in numbers of egg pods per square foot between field margins, between locations within margins, and between units within the locations, for 10 counties in north-central Montana in 1940; ½-square-foot sample units (degrees of freedom given in parentheses in column headings)*

County	Between field margins[1] (69)	Between locations within margins (1)	Interaction, field and location (69)	Between units within locations (560)
Blaine	1.12	0.82	[2]0.54	0.39
Cascade	.81	1.04	.42	.38
Chouteau	4.23	1.90	.87	1.02
Fergus	2.99	[2]5.32	1.22	1.33
Hill	8.62	[2]29.61	[1]5.07	3.29
Judith Basin	.13	.07	.04	.05
Liberty	1.19	[2]2.76	.56	.48
Pondera	6.16	7.20	3.98	4.13
Teton	11.33	.08	5.20	6.49
Toole	.86	.97	.46	.37

[1]Significant at 1-percent level of probability.
[2]Significant at 5-percent level of probability.

variations between field margins were highly significant for all 10 counties. Populations in the two halves of the margins were not significantly different in 7 of the 10 counties.

In general, the analyses show highly significant variations in the numbers of egg pods present in different fields and field margins. On the other hand, the variations between locations within fields or field margins had very little, if any, significance.

In order to make a more detailed comparison of the numbers of egg pods per square foot, 10 wheat-stubble fields in a single South Dakota locality were extensively sampled in 1942. Fifty pairs of ½-square-foot units from the same locations were taken in each. The pairs were taken from similar uniformly distributed locations in each field. The populations in the different fields ranged from 0.14 to 0.93 egg pod per square foot. Analysis of variance of the pairs is shown in table 5.

TABLE 5.—*Variance in numbers of egg pods per ½ square foot in one locality in South Dakota*

Areas sampled	Degrees of freedom	Mean square
Between fields	9	[1]1.16
Between locations within the same field	490	[2].26
Between units within the same location	500	.20

[1]Significant at 1-percent level of probability.
[2]Significant at 5-percent level of probability.

As was expected, the field-population differences were highly significant. Population variations due to locations within fields were significant, but the difference was not great and probably could not have been recognized, except for the large number of samples involved.

It has been pointed out that in some habitats, such as small-grain stubble, the number of egg pods approximated fairly consistently the average number in all the habitats (table 1). This fact, plus the fact that small grain is well distributed over the northern Great Plains, suggested the possibility that this habitat alone might be used in conducting the general survey. The experimental sampling both in grainfields and in other common types of habitat in two central Montana counties in 1939 seemed to indicate the reliability of such procedure.

Further evidence on this phase of the problem was obtained from the 10-county survey in north-central Montana in 1940. The average egg-pod populations for all fields, all margins, grainfields, and fields other than grain and their margins are given in table 6. Since field margins comprise

TABLE 6.—*Mean egg pod populations per square foot for fields, margins, weighted fields and margins, grain-stubble fields, and fields other than grain and their margins in 10 counties in north-central Montana in 1940*

County	All fields	All margins	All fields and margins	Grain-stubble fields	Fields other than grain and their margins
Blaine	0.59±0.07	0.83±0.08	0.60±0.06	0.40±0.03	0.82±0.12
Cascade	.52± .07	.65± .08	.53± .06	.41± .07	.68± .11
Chouteau	.46± .06	.90± .16	.48± .06	.36± .05	.87± .17
Fergus	.41± .06	.66± .14	.42± .05	.37± .08	.47± .07
Fergus[1]	1.80± .16	2.87± .77	1.85± .15	1.91± .18	1.40± .27
Hill	.40± .08	1.21± .22	.44± .07	.44± .08	.46± .18
Judith Basin	.10± .02	.08± .03	.10± .02	.11± .03	.08± .03
Liberty	.36± .05	.68± .08	.38± .04	.36± .05	.53± .14
Pondera	.77± .06	2.70± .19	.86± .05	.77± .06	1.01± .09
Teton	.58± .10	1.87± .26	.64± .09	.48± .09	1.20± .24
Toole	.41± .05	.51± .07	.41± .05	.41± .06	.41± .07

[1]1939 survey.

only about 5 percent of the total farm area, the averages for fields and for margins were given weights of 95 and 5, respectively, in obtaining

the means given in the columns "All fields and margins" and "Fields other than grain and their margins." It may be seen that the populations in field margins ran consistently higher than did those in the fields; however, the weighted means for both fields and margins were only slightly greater than those for the fields. A comparison of the weighted means for fields and margins with the means for the grain-stubble fields alone shows that the two groups of figures are close together. Possibly the most critical comparison is that of the average populations in small-grain stubble with those in the fields other than small grain and their margins, given in the last two columns. These sets of figures, although somewhat farther apart, still are close enough together to allow the grain-field figures to serve as a fairly reliable estimate of the general population. As a rule, grainfield populations appear the lighter. Standard errors of weighted averages are calculated as shown by Snedecor.[5]

Although these data substantiate the theory that the populations in grain stubble are a reliable index of the general population, there are reasons why it would not be advisable to sample grain stubble only. In the first place, small-grain fields, which comprise about 80 percent of the farmed acreage in the northern Great Plains, represent a preponderance of the agricultural area; therefore a general survey of this region, with stops prorated among the different habitats, would automatically be made up largely of small-grain fields. In the second place, the inclusion of the 20 percent of nongrain fields and field margins would provide for exploratory sampling to locate concentrations that otherwise might not be discovered.

SAMPLING AND SURVEY METHODS

In studying survey methods, the general nature of the population studied must be considered. If egg-pod distribution were entirely random, it would conform to the Poisson series,[6] and samples would reflect this distribution. In practice, a departure from the random or Poisson condition occurs, population being "bunched," and this departure is greater in dense infestations. In many sparse infestations the Poisson is nearly realized. In dense infestations variation is absolutely greater and proportionally less than in light ones.

The nearness to the Poisson, or random, distribution can be judged by the variance. In a true Poisson series the variance is equal to the mean. In the typical "bunched" condition the variance becomes greater than the mean. Under field conditions it is practically never found to be less than that of the Poisson. For this reason a greater degree of precision than that limited by the total number of sample units and a variance equal to the mean cannot be expected. An example may be drawn from the analysis of variance of the South Dakota data (table 5). In this rather sparse infestation the mean was 0.18, the variance between adjacent units was 0.20, and that between units well separated in the field was 0.26.

---

[5]See p. 3, footnote 4, (ch. 17).

[6]The Poisson series is explained by Snedecor's text (see p. 3, footnote 4). It is based on simple probability in population problems such as this, of the numbers of units with no egg pods, and with one, two, or more.

Number and Location of Field-Sample Units

Data presented above show that populations vary considerably more between fields than between units within fields, that there is a tendency to greater variation between distant units in a field than between adjacent units, but that in light populations neither type of variation is pronounced. Two adjacent units of ½ square foot each can be viewed as one 1-square-foot unit, thus affording a basis for comparing numbers and sizes of units. If two adjacent units were exactly similar, one would be as good as two, but if they differ greatly no special gain would come from spreading the sampling.

The South Dakota data (table 5) has been used in calculating expected standard error of different combinations, using established methods.[7] These calculations are given in table 7.

Table 7.—*The reliability of a grasshopper egg-pod survey based on different combinations of fields, field-sample locations, and units per field location* [1]

Fields sampled	Sample locations per field	Units per sample location	Standard error of mean egg pod population per unit[1]
Number	*Number*	*Number*	
20	5	1	0.052
20	5	2	.042
10	10	1	.057
10	5	2	.059
10	5	1	.074
10	1	10	.077
5	5	2	.084
10	1	5	.089

[1]Each field unit contained ½ square foot. To place the standard error on a 1-square-foot basis it must be doubled.

It can be seen that there is an advantage, but a very limited one, in spreading sampling within fields. Taking ½-square-foot units instead of 1-square-foot units, but doubling the number, will give only a slightly lower standard error. Taking the same number of units, but reducing their size by one-half, will increase the standard error considerably. To take all units in one place is going too far in the direction of "bunching" and increases the standard error markedly. In light infestations, such as the example just given, five 1-square-foot units are practically as good as twice the number of units half that size; but in denser infestations there would probably be more advantage in spreading within-field sampling. The five 1-square-foot units seem to offer a workable and sound combination. An increase in number of fields is more potent in reducing standard error than are increases in within-field sampling.

Distribution of Units Within Fields

The data in tables 2 and 3 show that location within the field has little systematic effect on the accuracy of the results. This fact indicates that sampling within the field need not follow an exact pattern that gives

[7]See p. 3, footnote 4 (sec. 17.8).

fixed representation to parts near the center and the edge. A less exact plan of taking the samples over a large part of the field—for example, an arc cutting well into the field and returning to the margin at another point—will do as well.

Sampling in the field margin may be carried out as the worker returns to the car. It would seem feasible to cut down marginal sampling. Table 6 shows that marginal populations had but little effect on weighted averages. While they are higher and more variable than field populations, the margins constitute only a small proportion of the total area in the northern Great Plains. Taking only two 1-square-foot units in the margin will not mean much sacrifice in accuracy.

All units should, of course, be located by some method that will preclude personal choice by the sampler. Randomness within fields is not essential in ordinary sampling, since the field itself is the primary sampling unit.

Number of Field Stops for a County or District

It has been shown that between-field variation is more important than that within field. In previous studies research workers have suggested a standard error of 0.125 egg pod per square foot as satisfactory. This level of precision is usually easier to reach in a low population than in a high one. It will be recalled that in the lightly infested area in South Dakota it could be achieved by moderate within-field sampling of 10 fields (table 4.).

Using the data from the 10-county survey of north-central Montana, the authors attempted to determine the number of field stops needed for each county and for the 10 counties grouped together as a district. The field margins were taken into account, being weighted on the basis of 95 percent for the fields, 5 percent for the margins. The allowable standard error of the mean was set at 0.125 egg pod per square foot. This concept of allowable standard error as a constant arithmetic figure is a useful one for the level of populations usually encountered. For sparse populations it would lead to using only a small number of units (table 8) and some minimum should be specified.

Table 8.—*Number of field stops necessary to survey each of 10 counties in Montana, allowing a standard error of 0.125 egg pod per square foot*

County	Approximate population mean	Stops needed to obtain a standard error of 0.125 in—	
		Fields only	Fields and margins
Judith Basin	0.1	3	3
Liberty	.4	10	9
Toole	.4	12	12
Fergus	.4	13	12
Hill	.4	28	27
Chouteau	.5	21	20
Cascade	.5	26	24
Blaine	.6	26	25
Teton	.6	37	35
Pondera	.8	16	15

It will be recalled that 70 fields per county were sampled and the fields sampled were prorated among the different habitats according to their acreages. Five pairs of ½-square-foot samples were taken in each field and 10 ½-square-foot samples along each margin. The survey data were treated by using variance between fields within each county. The results are shown in table 8.

Where the mean egg populations for the county ranged from 0.5 to 0.8 pod per square foot, from 15 to 35 field stops were required in order to obtain county means within the prescribed standard error in a survey by counties. The average number of fields needed was 20. For county means of about 0.4 egg pod per square foot, from 9 to 27 fields, or an average of 15, were needed. In South Dakota, where the mean population was about 0.2 egg pod per square foot, 10 fields were adequate. For Judith Basin County, Mont., with a mean population of 0.1, only 3 fields would have been necessary. Thus, it can be seen that the number of field stops needed per county for a survey on a county basis is largely affected by the population level. In low populations more than the calculated minimum just mentioned will be required for a representative survey.

Variation between county populations was not marked, but with the large number of fields used (70 per county) it appeared as highly significant. Analyzing with field means as units, the variance between counties was 2.14 and between fields within counties 0.29. It was found, however, that the number of fields necessary to represent the district with the required degree of precision was but a little greater than the number needed for a county having the same population level.

Moderate numbers of fields per county tended to give an unsatisfactory standard error for the county, and a standard error lower than needed for the mean of a group of similar counties. Less than 30 fields would have been required to give a standard error of 0.125 in the 10-county district considered (table 8); whereas for county units they total 182, and more would be required for representativeness in one or two of the counties. It is thus indicated that considerable economy might be effected by surveys of homogeneous groups of counties, rather than of individual counties as units.

Since the number of stops required for a given standard error increases with the density of population, a procedure with some elasticity in number of fields would seem desirable. A given minimum (say 8 or 10) sufficient to provide representativeness might be surveyed. If the standard deviation were calculated among the first 10, the number required for the desired standard error (s_x) of the mean could be tentatively determined by solving the well-known equation $s_x = s\sqrt{n}$ in which s equals standard deviation among field means and n equals the number of fields.

Distribution of Stops Within a County or District

The field stops are the real sample units and their distribution is of considerable importance. To obtain complete randomness of distribution is not practical. In a statistical sense randomness is used in order to give every unit in a population a chance to be represented in the sample. Applications of error formulas in a strict sense are based on this concept.

A rather systematic distribution of stops is better suited to many of the objectives of survey work. Three factors that must be provided for are as follows: (1) Representation of important sections and crop types, (2) plan of travel, (3) freedom from personal choice in selecting stops. In travel careful planning to obtain economy and coverage is required. In general, location of fields may be easily predetermined as at a given distance from some known point.

A systematic sampling procedure, such as suggested, will tend to give more accurate results than random sampling. At the same time, if error is calculated as if the sample were fully random, the error estimate will tend to be too high. When the systematic sample gives results closely equivalent to results from a random sample, as often occurs in this work, there will not be much inaccuracy in treating it as a random sample, and what inaccuracy there is will be on the conservative side. Some randomness may be arranged after the three essential factors have been provided for.

A restricted random-sampling plan, or stratified sampling, is often followed, in order to ensure that some units fall in each subarea of crop type. In such a plan the variance between types or subareas should be removed in calculating the sampling error, because the precision of such a survey is properly determined from the variance within the types or subareas, rather than between them. In this kind of sampling it may be difficult to arrange satisfactory stratification when the number of units is small. It may not be important, however, if areas are rather homogeneous, if one crop type is of outstanding importance, and if eggs are widely distributed. In some other grasshopper-survey problems it may be important.

If a large number of fields (25 or more) can be sampled, it may be possible to represent each important subdivision by several fields, preserving essential randomness in the location of these fields within the subdivisions. Randomness in location is easily achieved by some system of drawing numbers. Use of a group of similar counties as a unit may aid in attaining this objective and may ensure greater precision. If the number of fields is limited, it seems best to distribute them rather widely, select them by some objective method, and analyze the data as if the fields had been taken at random. The procedure described is not entirely correct, but it does not lead to serious mistakes.

Time Utilization of Sampling

A final evaluation of sampling plans may rest on the total possible cost. A plan giving the desired precision may be too expensive. Although sampling within fields is less important in reducing error than that between fields, it is less costly, and may give the best result in limiting the total expense of doing considerable work in a field, once the field is reached.

Biometric theory enables us to determine the best distribution of samples between and within fields if we have some preliminary estimates of limitations on total cost, cost of sampling fields and units within fields, and between-field and within-field variances. For purposes of analysis, these costs have been estimated in terms of time requirements. The total cost of planning, reaching a field, writing up results, and other necessary work has been estimated at 1 hour. The total allowable time

per county has been estimated from survey practice at 12 hours. The time required per ½-square-foot sample has been estimated at 2 minutes; or 4 minutes for taking a sample from the field and a corresponding sample from the margin. It is probably true that overhead costs per field would be reduced if the number of fields were increased; however, for this approximate approach we must assume fixed costs.

This application of biometric theory represents a departure from the practice of estimating the work required for a desirable standard error, such as 0.125 pod per square foot. It substitutes for this practice the finding of a work arrangement that will give as low a standard error as possible under given limitations. This standard error may not be as low as would be desired, but it will be the lowest possible under the stated conditions.

The variance for fields (V_f) over and above that expected from within-field variation is estimated by subtracting the mean square within fields from the mean square between fields and dividing by the number of units per field. It seems best for this calculation to pool within-field variances (for locations, interaction, and within locations), since they do not differ much and since the variance thus estimated will be that to be expected in general sampling. This procedure is illustrated by using data from table 3 for Cascade County. All within-field variance (weighted average) is estimated as 0.42; between-field variance is 0.93; $V_f=(0.93-0.42)/10=0.05$. V_w is simply within-field variance. By use of calculus, the best combination may be derived as follows: $k=\sqrt{(V_w \cdot CD)/(V_f \cdot C)}$ and $n=T/(CD+kC)$,[8] where k = number of half-square-foot units per field, n = number of fields, CD = overhead cost per field, C = cost per unit, and T = total cost. For Cascade County, if $T=720$ minutes, $CD=60$, $C=4$, we may calculate $k=\sqrt{(0.42\times60)/(0.05\times4)}$, or about 11. Then $n=720/[60+(11\times4)]$, or about 7. The best combination is thus indicated as 7 fields and 11 units per field. The variance of the county mean for this combination is estimated as $0.05/7+0.42/77$, or 0.0126. The standard error is $\sqrt{0.0126}$, or a little over 0.11; doubled to apply to a square foot basis it is 0.23. A practical demonstration by varying k may be made. If k is taken as 7 instead of 11, n will be 8; the variance of the mean will be $0.05/8+0.42/56$, or 0.0137. If k is 15, n will be 6, and variance will be $0.05/6+0.42/90$, or 0.0130. In either case the variance is increased, although a wide latitude in number of fields sampled and number of samples per field gives little change with these low variances.

Similar calculations including samples from the field margins give similar results. Their inclusion in this analysis evidently makes little difference unless the variance ratio departs sharply from that of field samples. Hence the influence of modifying marginal sampling is expressed as a reduction of the time required per field unit. In previous studies, the 2-minute period allowed for each field unit was doubled because an equal number of units was taken in the margin. The possibilities in taking fewer units in the margins and thereby reducing the time per within-field sample to 3 minutes per unit, as well as in allowing 16 hours

---

[8] If it is desired to estimate k and n for constant standard error and minimum cost, k will be estimated as above and n as $(K \cdot V_f+V_w)/(K \cdot V_m)$, where V_m is the square of the standard error of the mean.

instead of 12, have been tested. Approximate best combinations for these conditions and estimated standard errors have been worked out and are presented in table 9.

Table 9.—*Estimated best combinations for field sampling to determine grasshopper egg-pod distribution, with limited total cost, in north-central Montana*

County	Variance		Best combinations and resulting standard error per square foot					
			12-Hour total				16-hour total, (3 minutes per unit)	
	Between fields (V_f)	Within fields (V_w)	4 minutes per unit		3 minutes per unit			
Blaine	0.06	0.38	[1] 7 × 10	[2] 0.24	[1] 8 × 11	[2] 0.22	[1] 10 × 11	[2] 0.20
Cascade	.05	.42	7 × 11	.23	7 × 13	.21	10 × 13	.18
Chouteau	.05	.34	7 × 10	.22	8 × 12	.20	10 × 12	.18
Fergus	.02	.31	6 × 15	.17	6 × 18	.16	8 × 18	.14
Hill	.06	.44	7 × 10	.24	8 × 12	.22	10 × 12	.20
Judith Basin	.00+	.06	5 × 17	.07	6 × 20	.06	8 × 20	.05
Liberty	.02	.20	7 × 12	.15	7 × 14	.14	9 × 14	.12
Pondera	.02	.40	5 × 17	.19	6 × 20	.16	8 × 20	.14
Teton	.08	.62	7 × 11	.28	8 × 12	.26	10 × 12	.23
Toole	.05	.21	8 × 8	.20	8 × 9	.19	11 × 9	.16

[1] n (number of fields) × k (number of ½-square-foot units per field).
[2] Standard error per square foot for the combinations.

In the general survey in north-central Montana a frequent combination is 7 or 8 fields per county and 10 units (5 square feet) per field. This is seen (table 9) to be a fairly efficient combination when cost is considered. With low population and small differences among fields, as in Fergus and Judith Basin Counties, Mont., the taking of more unit samples per field from fewer fields would seem to be practicable and economical. With higher populations and great between-field variation, the sampling of a larger number of fields with fewer within-field units would be better. In any event, it is necessary to sample a fairly large number of fields to avoid risk of missing altogether some important local infestation.

DISCUSSION

The findings from these studies of egg-pod populations may be briefly summarized as follows: Variation between fields is definitely higher than that within fields. Variation between units within fields does not follow a definite location pattern. Although this variation is such as to make advisable the taking of unit samples from several well-separated locations in a field, little is gained by taking more than five samples. Randomness is not needed within fields, but there should be freedom from personal choice. The reduction of area per unit from 1 square foot to ½ square foot causes marked loss in precision unless the number is nearly doubled. Margins are higher and more variable in population than fields, but need not receive much consideration because of their comparatively small area.

In many instances from 15 to 30 fields are needed to obtain a satisfactory standard error of the area mean. For this purpose not many more stops will be required for a survey of several similar counties than would be needed for a single county. Distribution of stops over an area should provide for representation of principal subareas or types, should consider economy in travel, and should exclude personal choice. Randomness is desirable but may be hard to attain. It may be attained in part by sampling more fields per county or district.

Several suggestions for improvement in efficiency of sampling have been considered. The greatest contribution to accuracy would be made by increasing the number of fields sampled, but this procedure would be rather expensive. Restriction of within-field sampling to permit increase in the number of fields is not promising, because so much work would be needed to sample only a few more fields. Modification of within-field sampling is also unpromising; increase in amount and distribution of more than five 1-square-foot units does not give much gain in information, and more "bunching" results in considerable loss. Surveying grainfields alone does not seem to offer enough advantage to compensate for the reduction in information. Reduction in work in the margins seems well justified. Stratification, and also modification of number of fields in accordance with intensity of infestation, have been previously discussed and seem to have promise. Prediction of regional needs and transportation must be attacked before local distribution.

The idea of consolidating similar and contiguous counties into a single district seems very promising. If there could be some lessening of dependence on county lines, adequate estimates could probably be made from fewer fields, travel could be planned more efficiently, and more representative sampling of important crop types could be carried out. Estimates of bait and transportation for regional needs could readily be governed by such a survey, although the utilization of county machinery might still be required to ascertain local distribution. The development of district sampling would make it easier to stratify and modify the intensity of the survey. Under this plan the total number of fields would be larger than for any single county, and greater freedom should be possible in planning for their selection.

CONCLUSIONS

It is suggested that five 1-square-foot units be examined in each field and two similar units in its margin. For convenience, the field units may be located on an arc cutting well into the field, and the sampler may return along the margin. Units should be taken by a method that would ensure freedom from personal choice. Stops should be distributed with respect to crops and localities so that the samples will be as representative as possible within the limitations of time and accessibility. Randomness is desirable but not fully attainable. A restricted random plan is worthy of study. Probably not less than 10 field stops should be made in a county or a group of similar counties, and 15 to 30 stops would be better.

The data presented in this circular indicate that increased accuracy and some reduction in the cost of surveys could be realized by using, as the survey unit, a group of similar counties, instead of a single county; by varying the number of stops within a county or district, above a fixed minimum, according to population, as indicated by preliminary sampling; and by stratification of sampling to ensure representation of important environments.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1949—837288

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington 25, D.C.—Price 5 cents